

SRQuality: Uma Análise Comparativa de Técnicas de Detecção de Anomalias de Requisitos de Software

Fábio Rodrigues Pereira,¹ Paulo Afonso Parreira Júnior (Orientador)¹

Mestrado

¹Departamento de Ciência da Computação – Instituto de Ciências Exatas e Tecnológicas Universidade Federal de Lavras (UFLA)
Caixa Postal 3037 – 37200-000 – Lavras – MG – Brazil

Início: 07/2021 – Término (previsto): 07/2023 – Aprovação: 01/2022

fabio.pereira3@estudante.ufla.br, pauloa.junior@ufla.br

Abstract. *The lack of knowledge about the strengths and weaknesses of pre-existing approaches for requirements smells is a gap identified in the current literature. Therefore, this work aims to perform a comparative analysis of existing approaches for identifying requirements smells in order to understand their effectiveness (coverage and accuracy), as well as their strengths and weaknesses. Other than that, it is intended to specify recommendations for improvements to the tools analyzed.*

Resumo. *O desconhecimento dos pontos fortes e fracos das abordagens pre-existent sobre anomalias de requisitos é uma carência identificada na literatura vigente. Por conseguinte, este trabalho candidata-se a realizar uma análise comparativa das abordagens para identificação de anomalias de requisitos pre-existent com o intuito de se conhecer sua efetividade (cobertura e precisão), bem como seus pontos fortes e fracos. Além disso, pretende-se especificar recomendações de melhorias para as ferramentas analisadas.*

1. Introdução

A Engenharia de Requisitos (ER) é uma especificidade do domínio da Engenharia de *Software*, que consiste em um conjunto de métodos, técnicas e ferramentas para definição, análise, documentação e validação dos requisitos de um produto de *software* [Valente 2020]. Ou seja, a ER engloba uma série de atividades que, uma vez realizadas, resultarão em um documento textual com a descrição dos requisitos que o *software* deve contemplar. Esse documento é amplamente conhecido como Documento de Requisitos - DR [Chitchyan et al. 2006].

De acordo com Valente (2020), a ER pode colaborar diretamente para a qualidade do *software*, ao buscar alinhar o que está sendo especificado no DR com a factual expectativa do cliente. Todavia, a elaboração de um DR de qualidade não é uma tarefa simples. No geral, requisitos de *software* são descritos em linguagem natural, o que pode fazer com que contenham subjetividades e inconsistências, comprometendo sua compreensão e interpretação [Nascimento et al. 2018]. Isso, por sua vez, pode levar a implementações equivocadas e/ou defeituosas, frustrando as expectativas dos usuários, gerando retrabalho,

provocando o aumento dos custos de desenvolvimento e a desmotivação da equipe, dentre outros prejuízos.

O conceito de anomalias de requisitos (*requirements smells*) surge no contexto da busca por qualidade em um DR. Anomalias de requisitos são indicadores de que há “algo errado” com a descrição do requisito. Alguns exemplos de anomalias são termos comparativos, tal como “melhor que”, e superlativos, tais como “melhor desempenho” ou “menor tempo de resposta”, quando aplicados à especificação de um requisito de *software* [Femmer et al. 2017]. Esses termos são considerados anomalias, pois trazem subjetividade ao texto do DR, prejudicando sua compreensão e interpretação. Por exemplo, o significado de “menor tempo de resposta” para uma pessoa pode diferir do de outra, podendo levar a implementações que não satisfazem às reais expectativas dos *stakeholders*.

Diversos trabalhos têm sido realizados nos últimos anos, propondo métodos, técnicas e ferramentas com o objetivo de contribuir para a detecção de anomalias de requisitos e, conseqüentemente, para a qualidade do DR. De acordo com o Mapeamento Sistemático da Literatura (MSL) realizado por Nascimento et al. (2018), 41 estudos foram publicados sobre esse assunto no período de 2013 (quando o termo anomalias de requisitos foi introduzido pela primeira vez) até 2018. Contudo, pode-se notar que poucos desses trabalhos têm avaliado a efetividade das propostas preexistentes, comparando-as umas com as outras e procurando identificar seus pontos fortes e fracos [Nascimento et al. 2018]. A carência de estudos sobre essa temática resulta no desconhecimento dos pontos fortes e fracos das abordagens preexistentes, que poderiam servir de apoio para: (i) construir abordagens mais efetivas, isto é, com melhor cobertura e precisão quanto à identificação de anomalias de requisitos; e (ii) facilitar a tomada de decisão por parte de pesquisadores e profissionais da indústria quanto à escolha das abordagens que melhor atendam às suas necessidades.

Esta pesquisa tem objetivo duplo: (i) cobrir o *gap* da literatura quanto à análise da efetividade das abordagens preexistentes para identificação de anomalias de requisitos; e (ii) especificar recomendações de melhorias para as ferramentas analisadas. Para isso, pretende-se realizar uma análise comparativa das abordagens para identificação de anomalias de requisitos preexistentes, destacando a efetividade, bem como os pontos fortes e fracos de cada uma delas. Uma vez realizada essa análise, pretende-se especificar recomendações de melhorias com intuito de tornar a detecção de anomalias de requisitos mais efetiva e acurada. A priori, a análise proposta denomina-se *SRQuality*, que é um acrônimo para *Software Requirements Quality* (Qualidade em Requisitos de Software).

Para alcançar os objetivos propostos neste mestrado, o seguinte método de pesquisa deve ser executado: (i) atualizar o MSL realizado por Nascimento et al. (2018), buscando identificar as abordagens mais recentes para identificação de anomalias de requisitos; (ii) a partir dos resultados do MSL atualizado, selecionar as abordagens que serão comparativamente analisadas neste trabalho; (iii) analisar as abordagens selecionadas, identificando a efetividade, os pontos fortes e fracos de cada uma delas; e (iv) especificar recomendações de melhorias para as abordagens analisadas. Mais detalhes sobre o método de pesquisa descrito acima são apresentados na Seção 4 deste trabalho.

O restante deste trabalho está organizado da seguinte forma: na Seção 2 é apresentada uma breve fundamentação teórica sobre anomalias de requisitos e sua relação

com a qualidade de *software*. Na Seção 3 e 4, são apresentados, respectivamente, os trabalhos relacionados e o método de pesquisa a ser utilizado para atingir o(s) objetivo(s) proposto(s) neste mestrado. Na Seção 5, as informações sobre o estado atual desta pesquisa são apresentadas e, por fim, na Seção 6, são apresentados os resultados esperados e as considerações finais deste trabalho.

2. Fundamentação Teórica

A qualidade dos requisitos de um *software* é um dos principais fatores que contribui para o sucesso ou o fracasso do projeto, pois eles fazem parte de um documento essencial e crítico para as etapas posteriores do desenvolvimento, a saber, o Documento de Requisitos (DR) [Naeem et al. 2019]. Quanto mais cedo forem descobertos indícios de má qualidade no DR, menores serão os custos para sua correção e menor será o impacto na qualidade do *software*. De acordo com Nascimento et al. (2018), há na literatura, desde 2013, consenso sobre a importância de se detectar sintomas de má qualidade na especificação de requisitos; a esses sintomas deu-se o nome de Anomalia de Requisitos (*Requirements Smells*).

Para Femmer et al. (2017), as anomalias de requisitos são sintomas de defeitos na especificação de requisitos de um *software*, com uma localização concreta e um mecanismo de detecção. Em geral, tais anomalias são definidas com base em normas confeccionadas para auxiliar profissionais quanto à especificação de requisitos, tais como as normas IEEE 830 (1998) e ISO/IEEE 29148 (2018). Por terem uma localização concreta no DR (por exemplo, uma palavra ou uma frase), as anomalias de requisitos podem ser detectadas por meio de técnicas de Processamento de Linguagem Natural, tais como *POS tagging* (*Part-Of-Speech tagging*), análise morfológica, dicionários e lematização [Femmer et al. 2017].

3. Trabalhos Relacionados

No trabalho de Nascimento et al. (2018), foi conduzido um Mapeamento Sistemático da Literatura (MSL) sobre o tema anomalias de requisitos, contemplando 41 estudos entre 2013 a 2018. Esse trabalho é de grande relevância para o mestrado proposto, pois ele mapeia e classifica as principais abordagens da literatura para identificação de anomalias de requisitos. Contudo, os estudos analisados pelo MSL em questão foram publicados até 2018 e, por isso, o mesmo encontra-se desatualizado. Faz parte do método de pesquisa deste trabalho, realizar uma extensão do MSL proposto por Nascimento et al. (2018) - ver Seção 4.

No trabalho de Femmer et al. (2017), é feita a apresentação do conceito de anomalias de requisitos e, após isso, é apresentada uma abordagem para identificação de anomalias extraídas a partir da norma ISO/IEEE 29148 (2018). O principal diferencial do trabalho proposto neste mestrado, em relação ao trabalho de Femmer et al. (2017) e dos vários trabalhos destacados no MSL realizado por Nascimento et al. (2018) Nascimento et al. (2018), é a realização de uma análise comparativa das abordagens preexistentes para identificação de anomalias de requisitos.

O trabalho de Arendse e Lucassen (2016) é o que mais se aproxima da proposta deste mestrado. Neste trabalho, os autores realizam uma análise comparativa entre 3 (três) ferramentas adotadas para a identificação de anomalias em DR e apresentam um conjunto

de *guidelines* para confecção de uma nova ferramenta para essa finalidade. O foco do artigo estava em avaliar a efetividade das ferramentas quanto a dois aspectos específicos da qualidade de requisitos de *software*: ambiguidade (relaciona-se com o duplo sentido ou significado na escrita) e atomicidade (diz respeito à integridade da escrita). O presente trabalho, no entanto, propõe-se a avaliar a efetividade das ferramentas quanto as anomalias propostas por Femmer et al. (2017), que são o estado da arte no que diz respeito a anomalias de requisitos.

4. Método de Pesquisa

Nesta seção é apresentado o procedimento metodológico adotado para a realização deste trabalho. A Figura 1 ilustra os passos do método de pesquisa.

Atualizar MSL. Inicialmente, deve-se realizar uma atualização¹ do Mapeamento Sistemático da Literatura proposto por Nascimento et al. (2018), buscando identificar e catalogar: (i) os tipos de anomalias de requisitos que ocorrem e quais técnicas têm sido utilizadas para identificá-las (POS, dicionários); e (ii) o apoio ferramental existente para identificação de anomalias de requisitos e quais são as suas características (total ou semi-automático, *open-source* ou proprietária, multi-linguagem ou não, entre outros).

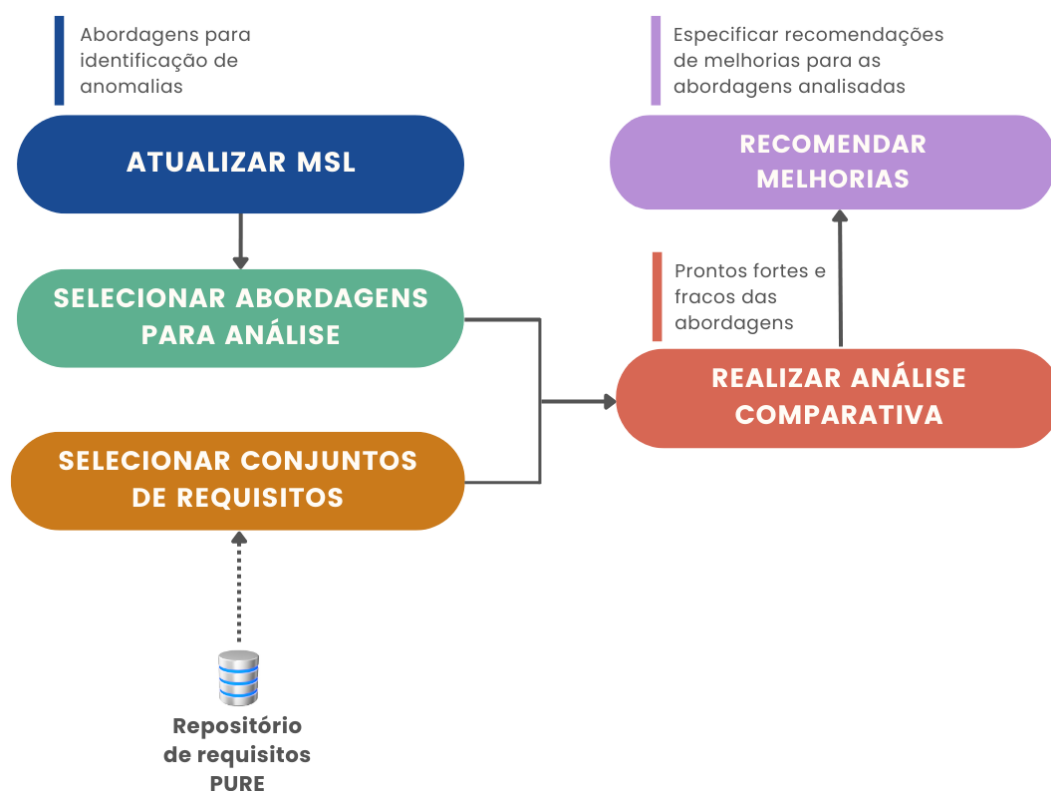


Figura 1. Esquematização do método de pesquisa proposto neste trabalho

Selecionar abordagens para análise. Com base no resultado da atualização do MSL realizado na primeira etapa do método de pesquisa, deve-se, então, selecionar as

¹A atualização de um MSL contribui com novos dados, evidências e resultados à partir de uma pesquisa realizada após sua data de publicação. [Nakagawa et al. 2017]

abordagens para identificação de anomalias de requisitos que serão comparados nesta pesquisa. Para isso, os critérios para seleção propostos por Arendse e Lucassen (2016) serão utilizados:

- **Automatização:** a abordagem deve ser totalmente automatizada por meio de um apoio computacional, evitando assim, que o conhecimento prévio do pesquisador influencie nos resultados;
- **Disponibilidade:** o apoio computacional deve estar disponível para *download* gratuitamente;
- **Documentação:** o artigo (ou página *Web*) que descreve o apoio computacional deve possuir informação clara e compreensível para sua instalação e uso; e
- **Formato de especificação dos requisitos:** o apoio computacional deve trabalhar com requisitos especificados em linguagem natural, que é o principal formato utilizado na construção de requisitos [Arendse and Lucassen 2016].

Selecionar conjuntos de requisitos. Deve-se selecionar uma coleção de Documentos de Requisitos a ser utilizado na análise comparativa das abordagens identificadas na etapa anterior. Essa coleção será extraída do *dataset* PURE (*A Dataset of Public Requirements Documents*), desenvolvido por Ferrari, Spagnolo e Gnesi (2017), e suas anomalias serão manualmente identificadas por especialistas na área de ER.

Realizar análise comparativa. Nesta etapa, deve-se realizar a comparação das abordagens selecionadas, por meio de seus apoios computacionais, utilizando a coleção de Documentos de Requisitos selecionado na etapa anterior. O objetivo é avaliar as ferramentas quanto à sua efetividade através das métricas de precisão (quantidade de anomalias apresentadas pela ferramenta que são, de fato, anomalias), e de cobertura (o percentual de anomalias identificadas pela ferramenta, em relação ao total de anomalias existentes).

Recomendar melhorias A partir dos resultados obtidos na etapa anterior, será especificado um conjunto de recomendações de melhorias, a fim de oferecer propostas de solução para os pontos fracos encontrados nas ferramentas e, se notáveis, recomendações para ampliar os ganhos dos pontos fortes.

5. Estado atual do Trabalho

Este mestrado encontra-se na fase de escrita da qualificação, a qual é um dos requisitos do Programa de Pós-graduação em Ciência da Computação da Universidade Federal de Lavras (PPGCC/UFLA). A qualificação deve ser entregue em até 12 (doze) meses após o início do primeiro semestre letivo do aluno regularmente inscrito no programa, o que ocorrerá em Setembro de 2022. Após a aprovação da qualificação, a próxima etapa será a defesa da dissertação de mestrado, que deve ocorrer em até 24 (vinte e quatro) meses após a matrícula do aluno no mestrado. Atualmente, o aluno encontra-se trabalhando na revisão bibliográfica sobre o assunto, realizando a extensão de MSL sobre anomalias de requisitos.

6. Resultados Esperados e Considerações Finais

Dada a importância de se ter um DR de qualidade e o entendimento que uma anomalia de requisito pode prejudicar as etapas seguintes do desenvolvimento do *software*, esse trabalho traz as seguintes contribuições científicas/tecnológicas: (i) atualização de um

MSL sobre anomalias de requisitos; (ii) cobertura de um *gap* de pesquisa sobre análise de abordagens para identificação de anomalias de requisitos; e (iii) elaboração de um conjunto de recomendações de melhorias para as ferramentas analisadas.

Referências

- Arendse, B. and Lucassen, G. (2016). Toward tool mashups: Comparing and combining nlp re tools. In *2016 IEEE 24th International Requirements Engineering Conference Workshops (REW)*, pages 26–31.
- Chitchyan, R., Sampaio, A., Rashid, A., and Rayson, P. (2006). In *A Tool Suite for Aspect-Oriented Requirements Engineering*, EA '06, page 19–26, New York, NY, USA. Association for Computing Machinery.
- Fabbri, S. C. P. F., Ferrari, F. C., and Camargo, K. G. (2014). A atividade de teste sob a perspectiva de qualidade de software. *Revista TIS*, 2(3):164–166.
- Femmer, H., Fernández, D. M., Wagner, S., and Eder, S. (2017). Rapid quality assurance with requirements smells. *Journal of Systems and Software*, 123:190–213.
- Ferrari, A., Spagnolo, G. O., and Gnesi, S. (2017). Pure: A dataset of public requirements documents. In *2017 IEEE 25th International Requirements Engineering Conference (RE)*, pages 502–505.
- IEEE (1998). Ieee recommended practice for software requirements specifications. *IEEE Std 830-1998*, pages 1–40.
- IEEE (2018). Iso/iec/ieee international standard - systems and software engineering – life cycle processes – requirements engineering - redline. *ISO/IEC/IEEE 29148:2018(E) - Redline*, pages 1–209.
- Naeem, A., Aslam, Z., and Shah, M. A. (2019). Analyzing quality of software requirements; a comparison study on nlp tools. In *2019 25th International Conference on Automation and Computing (ICAC)*, pages 1–6.
- Nakagawa, E. Y., Scannavino, K. R. F., Fabbri, S. C. P. F., and Ferrari, F. C. (2017). *Revisão Sistemática da Literatura em Engenharia de Software: Teoria e Prática*. Elsevier Brasil.
- Nascimento, R., Aranha, E., Kulesza, U., and Lucena, M. (2018). Requirements smells como indicadores de má qualidade na especificação de requisitos: Um mapeamento sistemático da literatura. In *WER*.
- Valente, M. T. (2020). *Engenharia de Software Moderna: Princípios e Práticas para Desenvolvimento de Software com Produtividade*. Independente.